



OPEN

SCB-YOLOv5: a lightweight intelligent detection model for athletes' normative movements

Qing Du¹, Lian Tang^{2,3✉} & Ya Li³

Intelligent detection of athlete behavior is beneficial for guiding sports instruction. Existing mature target detection algorithms provide significant support for this task. However, large-scale target detection algorithms often encounter more challenges in practical application scenarios. We propose SCB-YOLOv5, to detect standardized movements of gymnasts. First, the movements of aerobic athletes were captured, labeled using the labelling software, and utilized to establish the athlete normative behavior dataset, which was then enhanced by the dataset augmentation using Mosaic9. Then, we improved the YOLOv5 by (1) incorporating the structures of ShuffleNet V2 and convolutional block attention module to reconstruct the Backbone, effectively reducing the parameter size while maintaining network feature extraction capability; (2) adding a weighted bidirectional feature pyramid network into the multiscale feature fusion, to acquire precise channel and positional information through the global receptive field of feature maps. Finally, SCB-YOLOv5 was lighter by 56.9% than YOLOv5. The detection precision is 93.7%, with a recall of 99% and mAP value of 94.23%. This represents a 3.53% improvement compared to the original algorithm. Extensive experiments have verified that our method. SCB-YOLOv5 can meet the requirements for on-site athlete action detection. Our code and models are available at <https://github.com/qingDu1/SCB-YOLOv5>.

Keywords Digital sports, Action recognition, Deep learning, Attention module

The movements of aerobic gymnasts must be standardized, as this directly impacts their safety. Scientific and standardized movements can reduce or even eliminate the risk of injury. The new development stage of digital sports embodies the deep integration and interaction of “digital” and “sports”¹. Detecting and analyzing the behavior of aerobic athletes can help promote innovation in sports education, including actively promoting sports image recognition², 3D motion modeling analysis³, and live streaming⁴.

In recent years, human action recognition based on deep learning^{5–8} has found extensive applications in smart cities, industrial production, intelligent transportation systems, and other fields. The analysis of automated video content holds the potential to significantly advance monitoring capabilities, encompassing action recognition, target tracking, and pedestrian re-identification. Such advancements offer a practical approach to recognizing athletes' actions.

Precisely categorization and assessment of athlete behavior through visual data involves leveraging computer vision technology to predict action categories and evaluate action quality. Existing object detectors serve as important references for recognizing athlete behavior. However, deep neural models have a large number of model parameters, and their calculations are complex, imposing high demands on hardware computing capabilities, memory bandwidth, and data storage. This makes it costly to use them for practical sports education. Given the specialized and standardized nature of athletes' movements, further research is needed on existing recognition methods.

In summary, we have proposed a lightweight intelligent detection model: SCB-YOLOv5. This model is capable of analyzing human behavior in video images, identifying various types of actions, and responding promptly to specific circumstances, as illustrated Fig. 1.

We summarize the main contributions are as follows:

¹School of Resource Environment and Safety Engineering, University of South China, Hengyang 421001, China. ²School of Sports Science and Engineering, Hunan Institute of Engineering, Xiangtan 411100, China. ³School of Electrical Information Engineering, Hunan Institute of Engineering, Xiangtan 411100, China. ✉email: tanglian666@163.com



Figure 1. Our intelligent detection system for athlete normative behavior, including RGB cameras, clients, and alarms.

- We collect images of aerobic athletes, classify their actions anew, and establish the ANBD dataset for recognizing their behavior.
- In terms of model optimization, the SCB-YOLOv5 model is proposed. It features a more lightweight backbone and incorporates a weighted BiFPN to enhance the performance of the original model.
- We compared and analyzed comprehensive experiments with several approaches. The quality metrics of the target detectors were evaluated to validate the effectiveness of our approach.

Related works

Target detection algorithms

Target detection algorithms rely on convolutional operations⁹. Based on their framework structure, these algorithms can be categorized into two types: one-stage and two-stage models, as depicted in Fig. 2. One-stage algorithms generally prioritize high real-time performance and simplicity, often at the expense of detection accuracy. Conversely, two-stage algorithms utilize the regional proposal network (RPN) to generate suggestions and then employ a fully connected layer to produce category predictions and bounding boxes, resulting in higher detection accuracy. Despite having fewer network layers, some one-stage algorithms^{10–12} have recently outperformed two-stage networks in both accuracy and speed, and are widely employed in automated detection applications.

Actions recognition

Deep learning-based algorithms for recognizing athletes' normative actions, encompassing classification and detection. These capture video sequence images of athletes' movements and employ convolutional neural networks for model training to predict their actions. Numerous scholars have delved into human action recognition from a kinesiology perspective.

Zhang et al.²³ proposed a method that relies on multimodal sequence fitting to detect the behavior of college basketball players. This method combines motion patterns and visual motion features captured by cameras. Integrating global and local motion patterns can significantly improve the performance of group behavior recognition. Julian Fritsch et al.²⁴ introduced an intelligent detection algorithm for recognizing the post-scoring emotions of volleyball players, achieving a precision rate of 80.09%. Zhao et al.²⁵ focused on deep video analysis, extracting frame sequences as inputs for a 3D convolution-based deep neural network. This algorithm automatically captures spatio-temporal features of athlete behavior, thereby enhancing the accuracy of recognizing body movements.

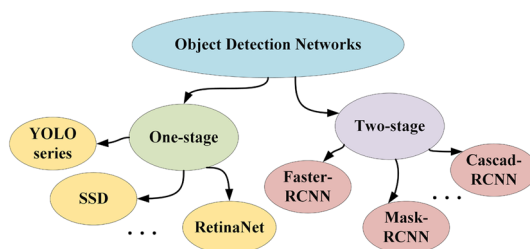


Figure 2. Existing target detection algorithms based on deep learning, where one-stage include YOLO series^{13–17}, SSD¹⁸, RetinaNet¹⁹, etc. and two-stage networks include Faster-RCNN²⁰, Mask-RCNN²¹, Cascad-RCNN²², etc.

Our dataset

Deep learning-based target detectors necessitate a substantial number of pre-labeled samples to enhance accuracy and generalization capability²⁶. We categorized movement classification based on fundamental body postures, basic techniques, and coordination. To create the athlete normative behaviors dataset (ANBD), we utilized pictures and video footage captured by members of a university aerobics team in Hunan Province, China. These videos were edited to extract one frame every 5 s, resulting in a collection of 2121 images showcasing various athletes in different scenes and angles.

The determination of whether a movement constitutes a standardized action relies on the varying amplitudes of the athlete's arm, elbow, leg, and other movements, as illustrated in Fig. 3. Among these, (a) depicts the standardized correct action, labeled “correct”; (b) illustrates the wrong hand action, labeled “wronghand”; (c) portrays the wrong leg action, labeled “wrongleg”.

Method

SCB-YOLOv5 model

All the images used in this research were obtained from the Hunan Institute of Engineering in Xiangtan, China, with 216 individuals, comprising 19 teachers and 197 students. All volunteers who participated in the photo shoots were informed about the data usage and provided consent for the research presented in this paper.

The overall structure of the SCB-YOLOv5 is derived from YOLOv5 and mainly consists of five components: input, Backbone, Neck, Head, and Predict. The specific structure is shown in Fig. 4. ShuffleNet v2²⁷ servers as the backbone to achieve a more lightweight design, integrating the CBAM at the base layer to capture additional feature information. The Neck component comprises BiFPN, which integrates semantic information from the deep network into the shallow network. Finally, the output predicts image features and generates the bounding box with the highest confidence based on the size of the target.

Mosaic-9

The mosaic data enhancement in YOLOv4¹⁴ randomly selects four images from the training set and combines their contents to create synthesized images directly used for training. This data augmentation method could enhance the ability of YOLOv4 to recognize objects in complex backgrounds. Therefore, we employ the Mosaic-9¹⁷ enhancement in YOLOv5, as illustrated in Fig. 5. Initially, a batch of images is randomly selected from the dataset, followed by the random selection of nine images from the extracted set. These pieces are then cut and stitched together to create a new image. This process is repeated batch size times (batch size refers to the number of images extracted from the dataset), resulting in an enhanced image of the specified batch size.

ShuffletNet V2

ShuffleNet V2²⁷ introduces new enhancements to the structure of ShuffleNet V1²⁸. The ShuffleNet V2 network structure is illustrated in Table 1. A 1×1 convolutional layer is incorporated to blend the features before the global average pooling. Efficient utilization of each stage enables an increase in feature channels and enhances network capacity. Notably, half of the feature channels in each block are directly transmitted to the subsequent one. This mechanism resembles feature reuse, akin to the concepts of DenseNet²⁹ and CondenseNet³⁰. Such a structure enables information communication between different channel groups and enhances reliability.

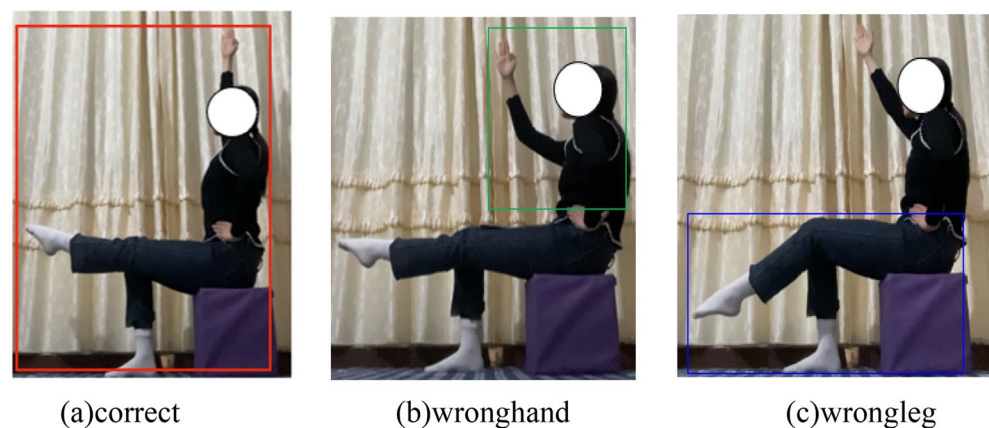


Figure 3. Classifying the movement behaviors of aerobics players, we categorized the common movements of the participants as follows: (a) the red box in the middle indicates that the athlete's arms and legs are straight, demonstrating standard movements; (b) the green box in the middle indicates that the athlete's hands are bent at the elbows, which is an incorrect movement that does not meet the normative requirements; and (c) the blue box in the middle indicates that the athlete's calves are bent, which is an incorrect movement that does not meet the standard movements.

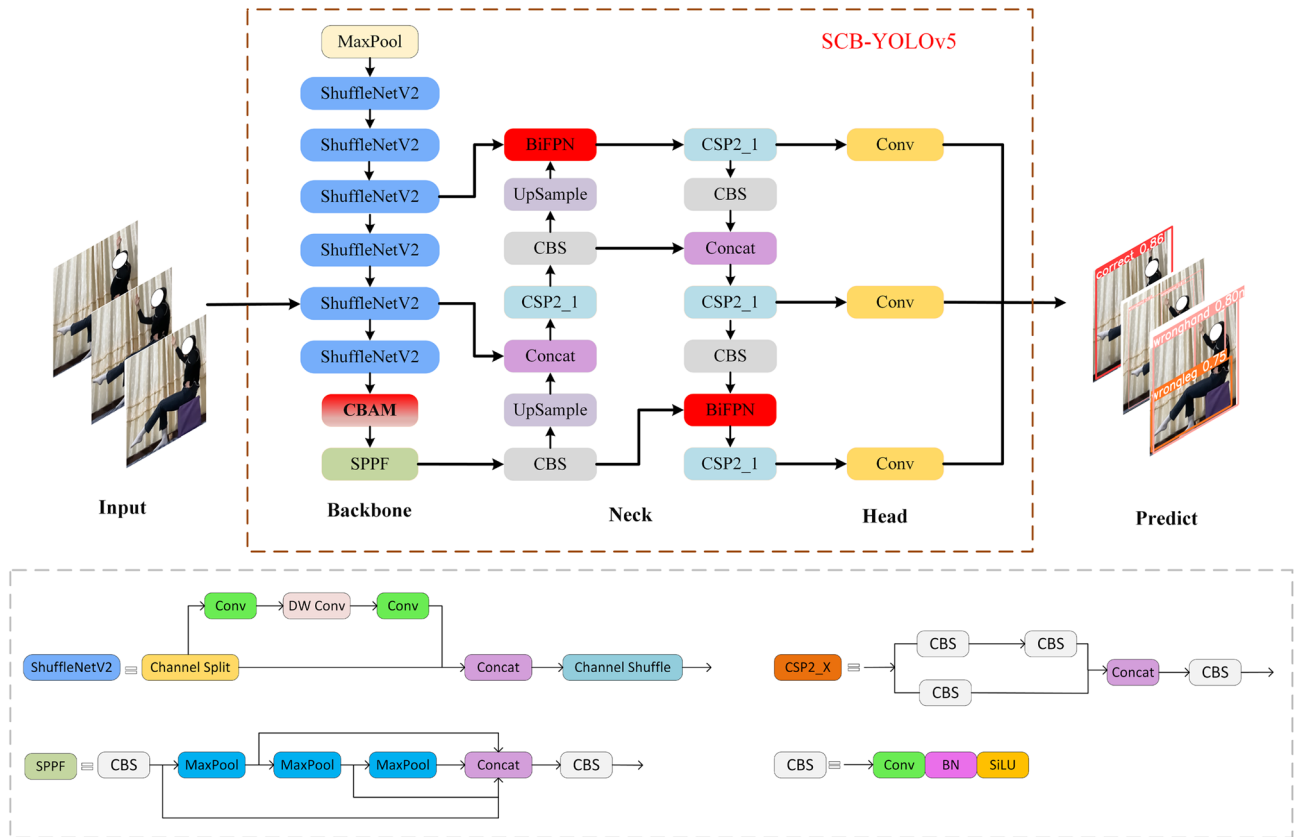


Figure 4. In the SCB-YOLOv5 structure, we primarily updated the backbone network and the Neck part of the YOLOv5.

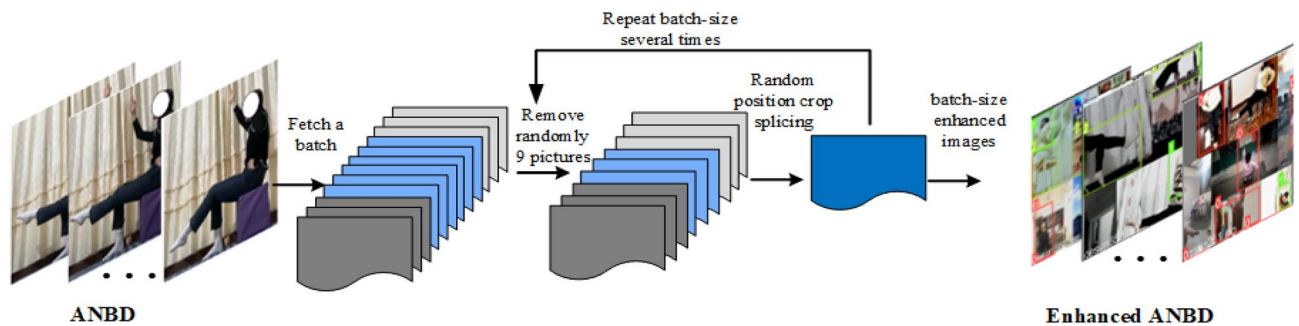


Figure 5. Mosaic-9. Using this strategy can effectively increase the amount of data, and our experiments verify that it has a positive effect on the subsequent model training.

Convolutional block attention module

Replacing the YOLOv5 backbone with the lightweight ShuffleNet V2²⁷ results in the loss of certain image feature information. To preserve more high-level semantic information, CBAM¹⁸ is added after the ShuffleNet V2, thereby directing more attention towards the significant aspects of the image. The CBAM attention mechanism is illustrated in Fig. 6.

CBAM comprises a channel attention module and a spatial attention module. It can effectively prioritize information crucial to the current task goal, enhancing the relevance of the extracted features from convolutional layers, capturing more comprehensive high-level semantic information, and improving target recognition. The calculation formula is as follows:

$$F' = M_C(F) \otimes F \tag{1}$$

$$F'' = M_S(F') \otimes F' \tag{2}$$

Layer	Out size	KSize	Stride	Repeat	Output channels			
					0.5 ×	1 ×	1.5 ×	2 ×
Image	224 × 224				3	3	3	3
Conv1 MaxPool	114 × 114 56 × 56	3 × 3 3 × 3	2 2	1	24	24	24	24
Stage2	28 × 28 28 × 28		2 1	1 3	48	116	176	244
Stage3	14 × 14 14 × 14		2 1	1 7	96	232	352	488
Stage4	7 × 7 7 × 7			1 3	192	464	704	976
Conv5	7 × 7	1 × 1	1	1	1024	1024	1024	2048
GlobalPool	1 × 1	7 × 7						

Table 1. ShuffleNet V2 structure, for each stage, its first block is required to be doubled, and the step size strips are all equal to 2.

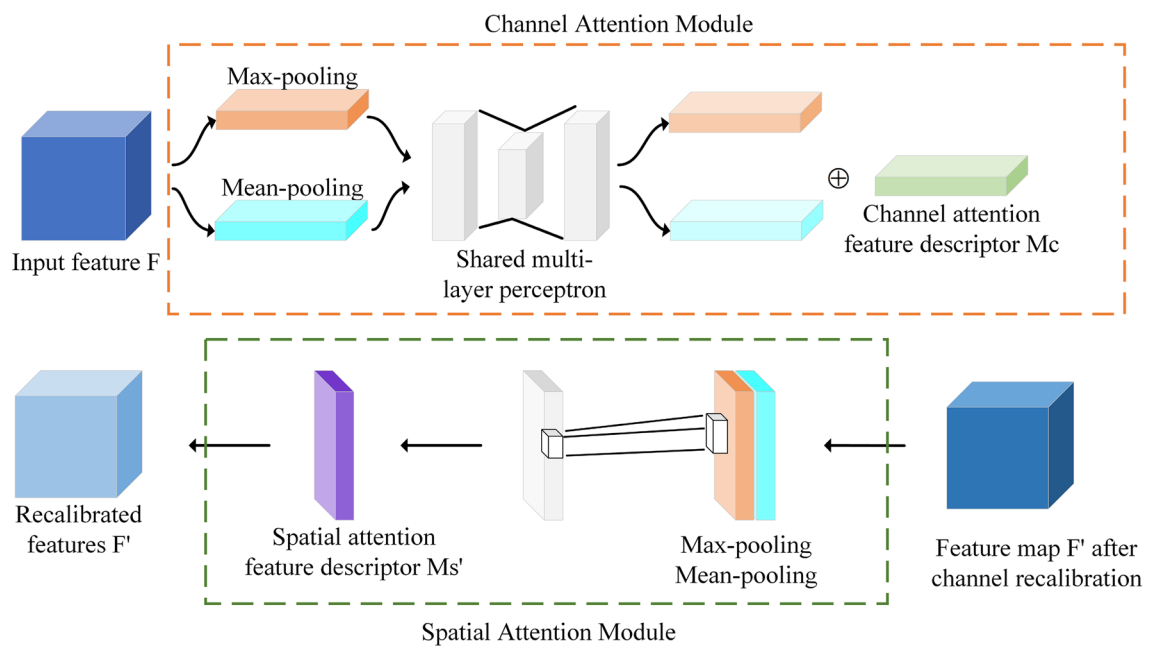


Figure 6. Structure of CBAM.

while, $F \in R^{C*H*W}$ is the input feature, $M_C \in R^{C*1*1}$ is the one-dimensional convolution of the channel attention module, $M_S \in R^{1*H*W}$ is the spatial attention module, F' is the output feature after passing through the channel attention module, and F'' is the final output feature.

Fusing the neck part of the BiFPN

The significance of analyzing images at multiple scales arises from the inherent complexity of images. Real-world scene images encompass a multitude of large and small target objects, each bearing diverse information such as size, position, color, and other attributes. Hence, relying solely on the bottom-up FPN pyramid structure has the potential to overlook information across different scales. To address this issue, Mingxing Tan et al.¹⁰ proposed BiFPN, a straightforward and efficient feature pyramid network, as depicted in Fig. 7.

The multi-scale feature fusion of BiFPN aims to aggregate features with different resolutions. Because the input features have different resolutions, BiFPN uses a band-weighted feature fusion method (Fast Normalized Fusion).

$$O = \sum_{i \in \epsilon + \sum_j w_j} \frac{w_i}{\epsilon + \sum_j w_j} \cdot I_i \tag{3}$$

where $\epsilon = 0.0001$ is used to avoid numerical instability, and w is a learned parameter, similar to an attention mechanism, used to distinguish the significance of various features in the feature fusion process.

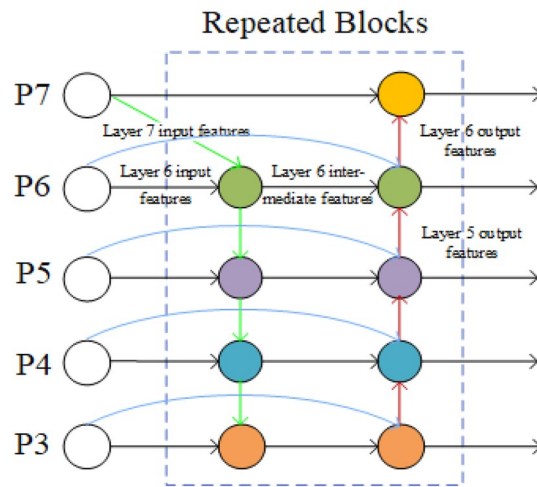


Fig. 7. BiFPN Node Diagram. The purple curve connects input nodes and output nodes within the same layer. The blue curve conveys semantic information of high-level features, while the red curve conveys positional information of bottom-level features.

Experimental results and discussion

Experiment settings

All our experiments were conducted using the Windows 10 system, utilizing the PyTorch deep learning framework. The processor employed is Intel(R) Core(TM) i5-10400F CPU @ 2.90 GHz, with 16 GB of RAM, and the GPU model is NVIDIA GeForce RTX 1650 graphics card with 4 GB of memory.

The dataset utilized for the experiments is ANBD, which consists of 2121 images as detailed in our dataset. The dataset is divided into training and validation sets in an 8:2 ratio. Model training encompasses 100 epochs, with a batch size set to 2, and the initial learning rate set to 0.01.

Model performance evaluation metrics

Average precision (AP) and mean average precision (mAP) are commonly used in target detection to evaluate the detection algorithms. The calculation formulas are presented in Eqs. (4) and (5):

$$AP = \int_0^1 PdR \quad (4)$$

$$mAP = \frac{\sum_{i=1}^k AP_i}{k} \quad (5)$$

where AP is the average precision of a single category, mAP is the mean of the AP values of all categories, F1 is the reconciled mean of P and R, P is the precision rate, R is the recall rate, and k is the number of detected categories.

AP is a common metric for evaluating the overall performance of a detector. However, excessive emphasis on labeled positive samples while pursuing AP can result in a high number of false detections. In practical evaluations, the F1-Score is employed as the evaluation criterion, offering a more balanced and effective measure of overall performance. Following 100 epochs of training, a relatively high confidence threshold (confidence = 0.5) is usually set to filter out a large number of false detection frames. Subsequently, the performance is analyzed using the F1-SCORE. As shown in Fig. 8.

Comparison with detectors results

Extensive experiments were conducted, including multi-group visual quality comparisons and image quality assessment, as shown in Table 1. (1) Among the original models, we found the YOLOv8 delivers the best detection



Figure 8. The F1-SCORE of detectors, where SCB-YOLOv5 has the highest F1-SCORE, while the other detectors exhibit less favorable detection results for the classification of wronghand, again confirming Table 2 above.

performance. However, it comes with a significantly larger number of model parameters compared to YOLOv5, albeit offering a comparable detection performance at a slightly higher cost. Therefore, the experiment aims to enhance YOLOv5. (2) SCB-YOLOv5 greatly reduces the network parameters and minimizes hardware computation. The mAP reaches 94.23%, which is 3.53% higher than YOLOv5.

Detection demos

The detection demos are shown in Fig. 9, corresponding to the data in Table 2. In the SSD algorithm figures a-ii, b-ii, c-ii, f-ii, and g-ii miss the detection of “wronghand” behavior. Figure a-iv has the highest sensitivity with a detection confidence of 100%. Conversely, the YOLOX and RetinaNet algorithms are less effective in detecting the “wrongleg” behavior. Figures b-iii and e-iii demonstrate the mission of “wrongleg”. Finally, SCB-YOLOv5 achieves accurate detection results.

In developing the SCB-YOLOv5, we meticulously documented the impact of each adjustment in the experiment, with a specific focus on the changes in Precision, Recall, and mAP. As shown in Fig. 10, the black line graph demonstrates that SCB-YOLOv5 outperforms the other methods in each performance metric after training stabilization.

Ablation study

To evaluate the effectiveness of the improved module on the SCB-YOLOv5, the ablation study was performed on our dataset. The results are shown in Table 3. After replacing the original backbone with the ShuffleNet V2 network, there was a decrease due to the reduced complexity of the model. This was done to maintain the detection performance of the model. Consider incorporating an attention mechanism after the ShuffleNet V2 backbone, and subsequently integrating BiFPN across multi-scale features to improve SCB-YOLOv5. Through extensive experiments, it has been proven that the adopted optimization strategy can enhance the accuracy of detecting. The mAP value has increased by 3.53 percentage points compared to the original.

Conclusion

In this study, we introduce a dataset for detecting the actions of aerobic athletes. A lightweight algorithm SCB-YOLOv5 is designed to recognize and regulate actions. To innovate the application of digital sports teaching processes.

The results of multiple sets of experiments show that the enhanced model has a more significant impact on recognizing athletes’ irregular hand and leg movements, outperforming other detectors. This finding holds major significance in promoting the sustainable and healthy development of “Internet + Education”.

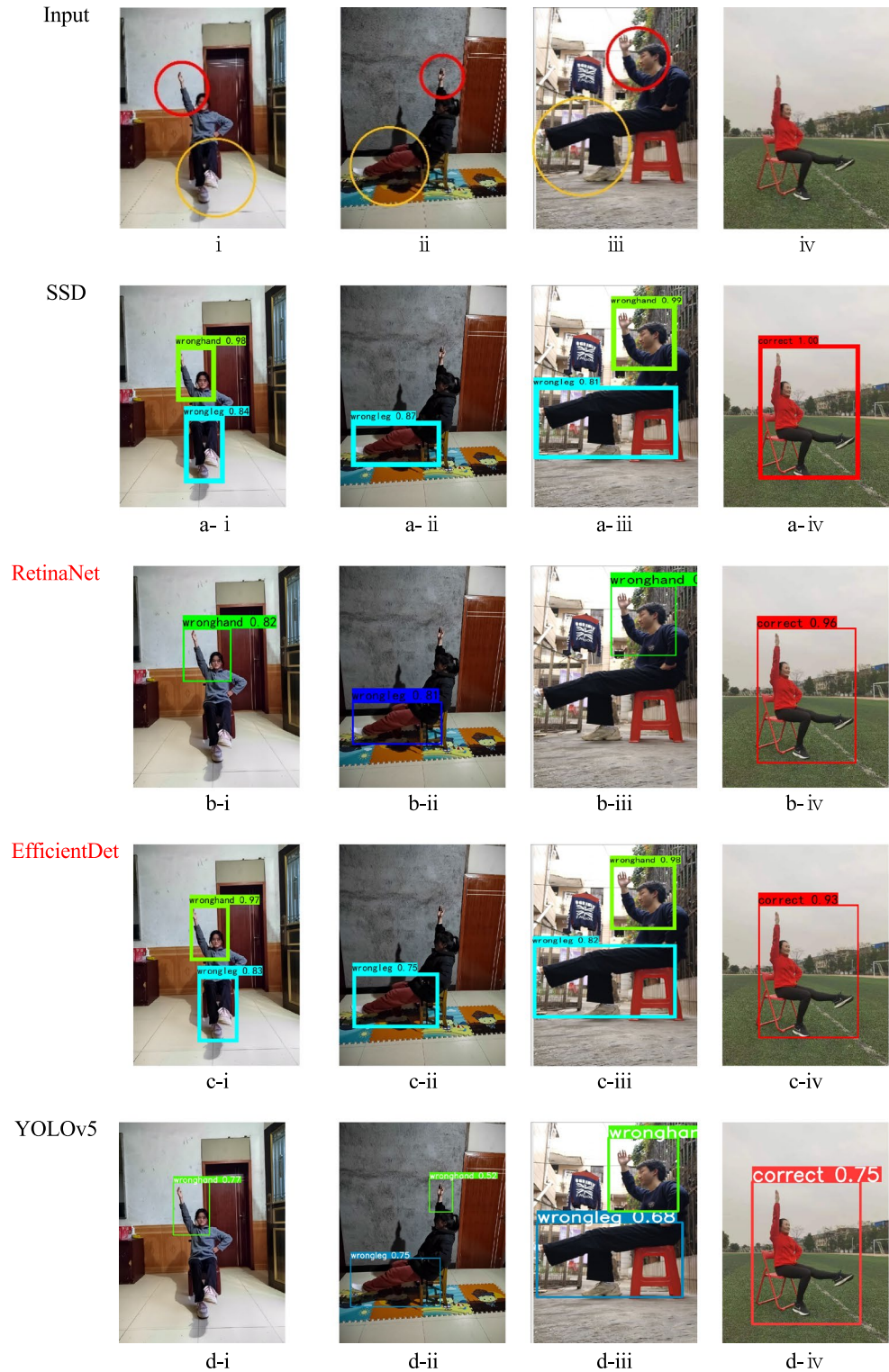


Figure 9. Detection demos with SSD, YOLOv5, YOLOX, YOLOv7, and YOLOv8 are shown in the figures i-iv depict diagrams of athlete behavior in various scenarios. Hand irregularities are marked in the red circles, foot irregularities in the yellow circles, and iv showcases correct behavior demonstrations. In figures a-i to f-iv, each detection box displays the detection confidence level, where higher values indicate greater confidence in the results of that detection.

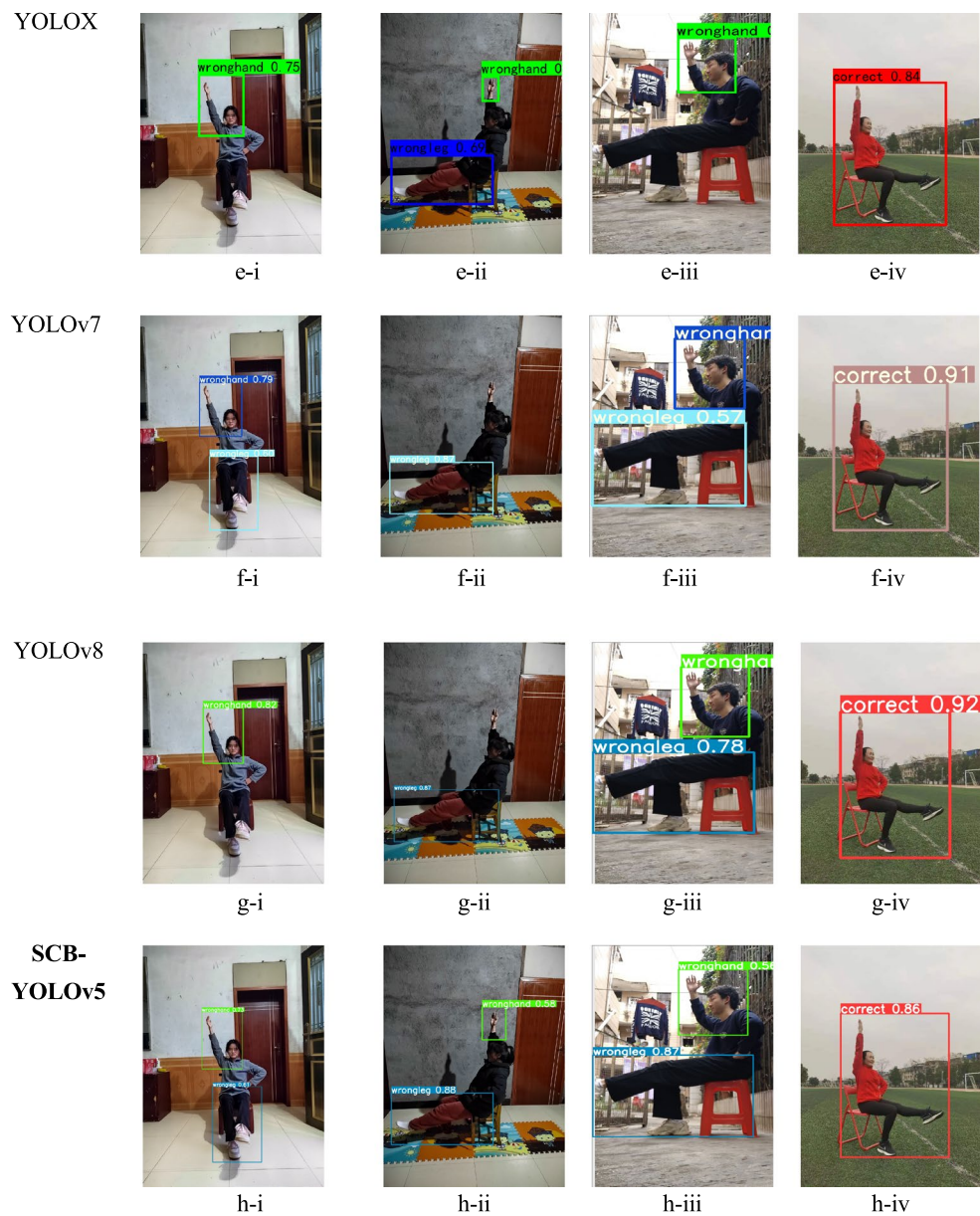


Figure 9. (continued)

Model	Backbone	Params (M)	AP			Recall	mAP
			correct	wronghand	wrongleg		
SSD	VGG		97.89	79.44	93.62	86.77	90.32
RetinaNet	ResNet50	34.0	95.90	81.53	92.03	81.06	89.82
EfficientDet	Efficientnet	6.6	91.64	80.83	82.30	89.18	84.92
YOLOv5	CSPDarknet53	7.2	96.15	84.04	91.92	89.16	90.70
YOLOX	Darknet-53	9.0	96.10	89.92	84.98	89.16	90.33
YOLOv7	E-ELAN	37.2	96.80	81.50	90.80	99.00	89.70
YOLOv8	CSPDarknet53	11.2	97.09	84.53	93.18	98.00	91.56
SCB-YOLOv5	ShuffleNet-v2	3.1	95.20	93.00	94.50	99.00	94.23

Table 2. Comparison of the detector results. The best experimental results are labeled in bold, and the second-best results are indicated in italic.

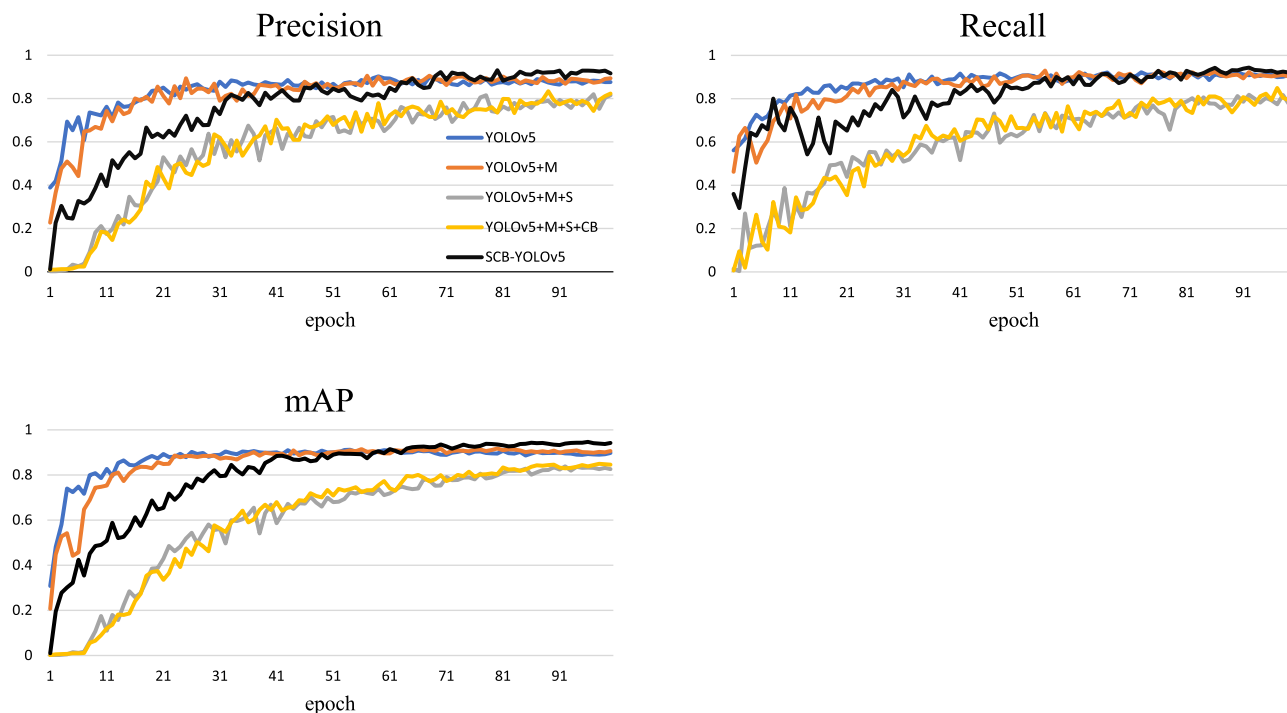


Figure 10. Line graph comparing the performance of each improvement experiment, where the horizontal coordinate is the 100epoch of the training of each algorithm model and the vertical coordinate is the index value.

YOLOv5	Mosaic-9	ShuffleNet-v2	CBAM	BiFPN	mAP
√					90.70
√	√				90.75
√	√	√			83.33
√	√	√	√		85.59
√	√	√	√	√	94.23

Table 3. Results of ablation experiments.

Data availability

Datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Code availability

Our code and models are available at <https://github.com/qingDu1/SCB-YOLOv5>. Our model code is open data. However, the data set used in this article is not available, and the authors can be contacted if necessary.

Received: 12 December 2023; Accepted: 8 April 2024

Published online: 14 April 2024

References

- Al-Emran, M., Malik, S. I. & Al-Kabi, M. N. A survey of Internet of Things (IoT) in education: Opportunities and challenges. In *Toward Social Internet of Things (SloT): Enabling Technologies, Architectures and Applications* (eds Hassaniien, A. E. et al.) 197–209 (Springer, Cham, 2020).
- Li, G. & Zhang, C. Automatic detection technology of sports athletes based on image recognition technology. *EURASIP J. Image Video Process.* **2019**, 1–9 (2019).
- Ghosh, P., Song, J., Aksan, E., Hilliges, O. Learning human motion models for long-term predictions. In *Proceedings of International Conference on 3D Vision*, 458–466 (IEEE, 2017).
- Levallet, N. et al. Enhancing the fan experience at live sporting events: The case of stadium Wi-Fi. *Case Stud. Sport Manag.* **8**(1), 6–12 (2019).
- Chen, D. D. Image recognition of sports athletes' high-intensity sports injuries based on binocular stereo vision. *Comput. Intell. Neurosci.* **2022**, 4322597–4322597 (2022).
- Batty, M. Big data, smart cities and city planning. *Dialogues Hum. Geogr.* **3**(3), 274–279 (2013).

7. Baines, T., Lightfoot, H., Smart, P. & Fletcher, S. Servitization of manufacture: Exploring the deployment and skills of people critical to the delivery of advanced services. *J. Manuf. Technol. Manag.* **24**(4), 637–646 (2013).
8. Zhu, K., Wang, R., Zhao, Q., Cheng, J. & Tao, D. A cuboid CNN model with an attention mechanism for skeleton-based action recognition. *IEEE Trans. Multimed.* **22**(11), 2977–2989 (2020).
9. Cao, D., Chen, Z. & Gao, L. An improved object detection algorithm based on multi-scaled and deformable convolutional neural networks. *Hum. Cent. Comput. Info.* **10**(1), 1–22 (2020).
10. Tan, M., Pang, R., Le, Q.V. Efficientdet: Scalable and efficient object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA*, 10781–10790 (2020).
11. Liu, Z., Lin, Y., Cao, Y., Hu, H., Zhang, P., Lin, S., Guo, B. Swin transformer: Hierarchical vision transformer using shifted windows. In *Proceedings of the IEEE/CVF International Conference on Computer Vision, Nashville, USA*, 10012–10022 (2021).
12. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirill, A., Zayko, S. End-to-end object detection with transformers. In *European Conference on Computer Vision*, 213–229 (Glasgow, 2020).
13. Redmon, J., Divvala, S., Girshick, R., Farhadi, A. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA*, 779–788 (2016).
14. Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M. Yolov4: Optimal speed and accuracy of object detection. In *Computing Research Repository*. <https://arxiv.org/abs/2004.10934> (2020).
15. Ge Z, Liu S, Wang F, et al. Yolox: Exceeding yolo series in 2021. Preprint at [arXiv:2107.08430](https://arxiv.org/abs/2107.08430) (2021).
16. Li C, Li L, Jiang H, et al. YOLOv6: A single-stage object detection framework for industrial applications. Preprint at [arXiv:2209.02976](https://arxiv.org/abs/2209.02976) (2022).
17. Wang C.Y., Bochkovskiy A., Liao H.Y.M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. Preprint at [arXiv:2207.02696](https://arxiv.org/abs/2207.02696) (2022).
18. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y. and Berg, A.C., Ssd: Single shot multibox detector. In *Proceedings of Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands* (2016).
19. Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P. Focal loss for dense object detection. In *Proc. IEEE International Conference on Computer Vision, Honolulu, USA*, 2980–2988 (2017).
20. Ren, S., He, K., Girshick, R. & Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**, 1 (2015).
21. He K., Gkioxari G, Dollár P, et al. Mask r-cnn. In *Proc. IEEE International Conference on Computer Vision*, 2961–2969 (2017).
22. Cai, Z., Vasconcelos, N., Cascade R-CNN: Delving into high quality object detection. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 6154–6162 (Salt Lake City, 2018).
23. Zhang, L. Behaviour detection and recognition of college basketball players based on multimodal sequence matching and deep neural networks. *Comput. Intell. Neurosci.* <https://doi.org/10.1155/2022/7599685> (2022).
24. Fritsch, J., Ebert, S. & Jekauc, D. The recognition of affective states associated with players' non-verbal behavior in volleyball. *Psychol. Sport Exerc.* **64**, 102329 (2023).
25. Zhao, X. P. Research on athlete behavior recognition technology in sports teaching video based on deep neural network. *Comput. Intell. Neurosci.* <https://doi.org/10.1155/2022/7260894> (2022).
26. Heffington, C., Park, B. B. & Williams, L. K. The “most important problem” dataset (MIPD): A new dataset on American issue importance. *Confl. Manag. Peace Sci.* **36**(3), 312–335 (2019).
27. Ma, N., Zhang, X; Zheng, H.T., Sun, J. Shufflenet v2: Practical guidelines for efficient CNN architecture design, 116–131. <https://arxiv.org/abs/1807.11164> [cs.CV] (2018).
28. Zhang, X., Zhou, X., Lin, M., Sun, J. Shufflenet: An extremely efficient convolutional neural network for mobile devices, 6848–6856. <https://arxiv.org/abs/1707.01083> [cs.CV] (2018).
29. Iandola, F., Moskewicz, M., Karayev, S., Girshick, R., Darrell, T., Keutzer, K. Densenet: Implementing efficient convnet descriptor pyramids. <https://arxiv.org/abs/1404.1869> [cs.CV] (2014).
30. Huang, G., Liu, S., Van der Maaten, L., Weinberger, K.Q. Densenet: An efficient densenet using learned group convolutions, 2752–2761. <https://arxiv.org/abs/1711.09224> [cs.CV] (2018).

Author contributions

Q.D.: conceptualization, methodology, software, writing—original draft; L.T.: conceptualization, writing—review and editing; Y.L.: data curation, writing—original draft.

Funding

This work was supported in part by the Research Foundation of Hunan Province Innovation Foundation for Postgraduate, China Grant QL20230233.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to L.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024